Use Case 13	**Automated Coding of Pathology Reports (caBIG/ U Pitt)**
Case:	Automated coding and extraction of concepts and concept relationships from Pathology Reports.
Background:	The University of Pittsburg Cancer Center (UPCC) is building an Information Extraction (IE) System that will be used to detect the occurrence of concepts and inter-concept relationships in Pathology Reports (PR). The IE will automatically code and classify the PR on the basis of concepts and relationships found in the PR. Currently, UPCC is interested in using the NCI Thesaurus as the ontology resource used by the IE System.
Requirements:	The NCI Thesaurus must provide comprehensive coverage of findings as well as diseases. Definitions of high quality are needed for all finding and disease concepts. Findings that are defining with respect to classifying disease concepts as subtypes of a broader disease concept must be differentiated from findings that are not defining with respect to sub-typing diseases. The clinical significance of findings to diseases must be expressed. Non-definitional relationships have prognostic value. Definitional and non-definitional relationships must be differentiated. Stage and Grade of disease concepts for each category of disease must be available in Thesaurus. The relationship of disease, pre-disease and benign conditions to cytogenetic and molecular abnormalities must be specified. A consistent approach must be followed across disease, pre-disease and benign concepts. Coverage of immunohistochemistry concepts must be expanded in order to be adequate for parsing pathology reports. Relationships between immunohistochemistry and diseases must also be defined. The version number of the terminology must be accessible to software interacting with the terminology server.

	Descriptions of planned enhancements to the terminology, especially with respect to semantic relationships, must be available to terminology users.
Solution:	NCI Thesaurus provides coverage of diseases, findings, molecular abnormalities, anatomy, and abnormal cells related to cancer science. EVS will review and augment content coverage and add concepts as needed. EVS is developing professional and lay definitions for all NCI Thesaurus concepts. For example, disease definitions are currently available for major cancers and non-cancerous neoplasms. Definitions are being created for less frequent/rare neoplasms, and for newly introduced terms. Findings that are defining with respect to a disease are indicated by the role Disease_has_Finding; other findings are indicated with the role Disease_may_have_Finding. Findings are organized into distinct categories, such as test results; by anatomic site/system, by cause, by morphology, etc. The clinical significance of a finding (e.g. aggressive clinical course, favorable clinical outcome) is related to the morphologic or laboratory finding associated with the particular outcome using a "role group." NCI Thesaurus disease modeling includes information on stage and grade using the roles Disease_is_Stage and Disease_is_Grade. Stage and Grade concepts are organized by coding system. Stage and grade concepts from different systems are associated as being equivalent where appropriate. NCI Thesaurus classifies neoplasms as benign, precancerous or malignant, and links them to related molecular and cytogenetic abnormalities using the roles Disease_ has/may have_Molecular_Abnormality and Disease_has/may have_Cytogenetic_Abnormality. Addition of immunohistochemical findings is a planned development project for NCI Thesaurus, in order to further define diseases. Currently a few lymphomas have been defined immunohistochemically according to their CD30 and CD15 positivity.

	Terminology version number is available to software in the NCI Vocabulary Server. A version tag will be added to the OWL version of the NCI Thesaurus data. A description of the current semantics of the NCI Thesaurus is available at the EVS download site. Planned or proposed enhancements will be added to the material available on the site.
Actions:	Continue UPCC-EVS collaboration. UPCC and other Pathology Informatics needs will be reflected in EVS work planning with respect to definition development and concept coverage extensions. UPCC will be asked to contribute definitions, lists of needed concepts, and to review pre-release EVS terminology products.

.